AD609791

MEMORANDUM
RM-4323-PR
NOVEMBER 1964

ASSESSING THE IMPACT OF WEAPON SYSTEM MODIFICATIONS BY FIELD STUDY: A FEASIBILITY DEMONSTRATION

Anders Sweetland

PREPARED FOR:

UNITED STATES AIR FORCE PROJECT RAND

The RAND Corporation
SANTA MONICA • CALIFORNIA

MEMORANDUM
RM-4323-PR
NOVEMBER 1964

ASSESSING THE IMPACT OF WEAPON SYSTEM MODIFICATIONS BY FIELD STUDY: A FEASIBILITY DEMONSTRATION

Anders Sweetland

This research is sponsored by the United States Air Force under Project RAND—Contract No. AF 49(638)-700 monitored by the Directorate of Development Plans, Deputy Chief of Staff, Research and Development, Hq USAF. Views or conclusions contained in this Memorandum should not be interpreted as representing the official opinion or policy of the United States Air Force.

The RAND Corporation
1700 MAIN ST • SANTA MONICA • CALIFORNIA • 90406

PREFACE

This is one of several RAND studies exploring ways to use computers for data analysis in weapon system management. This particular investigation is a product of the fourth project of our Logistics Simulation Laboratory (LP-IV), and shows how a weapon modification may be field tested early in the conversion to determine how it is affecting maintenance and operational programs.

The Memorandum should interest people in the Maintenance Engineering Directorates at both Hq USAF and Command levels who want to measure the cost and effectiveness of a modification program. Others who may find this research useful are those studying new methods for analyzing weapon system data.

For another study of the same general problem, interested readers may wish to consult K. L. Deavers, Product Improvement Candidate Selection and Management Information, The RAND Corporation, RM-4324-PR, November 1964.

SUMMARY

For a number of years, The RAND Corporation has been studying ways to increase the productivity of the Air Force maintenance management system. These studies have suggested adding clock-hours and status data to the current data system* (AFM 66-1, Maintenance Management), and have proposed the use of computers in maintenance management and analysis.** These additions have made practicable analyses of maintenance data that were heretofore deemed impossible. This Memorandum continues the previous line of inquiry. It demonstrates how these adjuncts to the present system may be used to assess the cost and effectiveness of engineering modifications.

We first describe the procedures using two examples: an engine modification of T-38A aircraft, and a fire-control system modification of F-101B aircraft. The two procedures are then contrasted to show how the combination of methods can provide the most meaningful set of data.

Next we show how the method of recording and analyzing may also be used to provide a history of the installation-checkout process during this vital period. It is worth comment that information of this kind -- elapsed time for completion, job standards, test-flights required, kinds of defects discovered, bit and piece consumption and man-hour utilization -- is the most difficult to obtain during the confusion of a heavy modification program.

Finally, we briefly discuss some factors that would be involved to complete the package described. The major addition to the present method would be the inclusion of mission data in a manner that could be related to the maintenance data.

---

*C. F. Bell, Jr., and T. C. Smith, The Oxnard Base Maintenance Management Improvement Program, The RAND Corporation, RM-3370-PR, November 1962.

**A. Sweetland, The Use of Computers in Air Force Maintenance Management and Analysis, The RAND Corporation, RM-4228-PR, October 1964.

ACKNOWLEDGMENTS

I am indebted to the following people for their helpful suggestions: Major E. D. Moore and Senior Master Sergeant A. Barrowman, ATC; Dr. L. Knudsen and Captain J. Abell, ADC; and Captain C. E. Mercer, ADC. Major J. Austin, ADC, helped write the recovery program used in the analysis.

CONTENTS

I. INTRODUCTION

Because of changing threat assessment and because of wear-out factors, a weapon cannot be tested under all circumstances before it comes into the operational inventory. As experience with the weapon reveals the need for improvements, they are developed, contracted for, and scheduled into the system. This is the weapon modification program.

Modifications are of three general types: (1) the basic repair, concerned with the prevention of wear-out; (2) the reliability-maintainability improvement modification, most often concerned with high-failure components; and (3) the performance improvement modification, concerned with improving mission capability or improving safety-of-flight. This Memorandum deals mainly with the latter two types.

It is difficult to estimate exactly what a modification program costs, but it can be considerable. For example, in 1962, Oxnard Air Force Base averaged 728 aircraft hours per month in Time Compliance Technical Orders (TCTO's) as reported by AFR 65-110A procedures. This is the equivalent of one aircraft in continuous TCTO status, which represents a loss in operational capability. Also, during the modification program described later, Oxnard was excused from active air defense for two months. This represents a severe loss in operational capability.

In terms of man-hour costs, the amounts can become substantial. For example, at one time, Air Defense Command (ADC) requested an average of 50 man-hours a month per possessed aircraft for modification purposes. Hence a squadron of 20 aircraft would average 2000 man-hours each month.

That the modification program is costly is apparent. The way cost is defined, of course, is most germane. Although there are many ways to define cost, the definition used here includes three factors:

1. The dollar cost of kits and engineering (including technical manuals.

2. The man-hour costs for installation and checkout, with civilian and military personnel kept separate.

3. The mission inventory loss, i.e., some representation of the reduction of mission effectiveness due to non-availability of the weapon during modification. This is generally expressed in two ways

   a. as a whole number, i.e., aircraft hours, and
   b. as a ratio of this number to weapon inventory, i.e., aircraft hours lost to tech-order programs divided by possessed aircraft hours.

To set our general dimensions right, let us briefly note the costs of two recent engineering modification programs. In 1963, Air Training Command modified the power plants of all T-38A aircraft. The cost of installation kits and engineering was $1.7 million. The average military man-hours used for installation and checkout was 275 to 300 per aircraft. In addition, a private firm was paid $15 million to assist in the installation. No data were kept on the loss of the weapon from the operational inventory.

During 1963, Air Defense Command did a fire control modification of F-101's. The engineering and kit cost was $21.9 million. No man-hour figures were available on the installation, which was done at the depot. The alignment and checkout was done in the field, and averaged 450 man-hours per aircraft (of this, two-thirds was flight-line maintenance and one-third was bench repair). The aircraft inventory loss was approximately five weeks per aircraft: three weeks for installation; one week for check-out and alignment; and one week for problems of scheduling, NØRS, and transportation.

Of course, not all modifications are of such magnitude. The majority are much more pedestrian in nature: replacing diodes to get better peak inverse voltage tolerances, rewiring to reduce high-voltage arcing, fusing circuits, increasing ventilation, etc. Although these are far less dramatic, each involves engineering and installation costs, and each installation generally results in some loss to the mission inventory. The accumulation of these small items typically adds up to a formidable total.

In costing a modification program, then, there are <u>at least</u> three areas that should be considered: the part's dollar cost, the man-hour cost, and the loss-of-mission cost. The assessment of costs, however, is only part of the task. A modification is installed to obtain one,

or more, specific objectives. The critical question becomes: To what extent did the modification accomplish its stated objective? This brings up the issue of how to measure effectiveness, the concern of a major part of this Memorandum.

Since most modifications represent an effort to improve either equipment reliability or maintainability, the major part of the effort, to date, has been devoted to obtaining measures in the reliability-maintainability area. More specifically, the major effort has been directed toward getting break- and recovery-rate distributions, i.e., both the means and the variances. Knowledge of the variances is necessary if we are to make definite statements about the impact of the modification on effectiveness. This point will become apparent in the following discussion.

II. THE SORTIE AS THE UNIT MEASURE

Early in the development of a package of computer programs to be used in analyzing maintenance data, it became evident that the ability to isolate data by the specific sortie was the sine qua non of any analysis in depth.* A specific sortie is one flown on a specific day and hour by a specific tail number. The data we wanted to isolate by sortie involved determining: the kind of mission flown, the systems and subsystems that required unscheduled maintenance, and the time and man-hours consumed in maintenance.

Specific sortie analysis contrasts with general sortie analysis in that the latter deals with aggregates (generally monthly totals). There is a tremendous difference in the amount of information each method yields. Take a typical example -- the unscheduled man-hours per sortie cost. In the case of the generalized sortie, the total unscheduled man-hour cost is determined and then divided by the number of sorties flown. The result is the average number of man-hours per sortie, a single quantity.

In analyzing the specific sortie, the unscheduled man-hour cost is determined by processing the individual sortie measures. Hence, it is possible to get not only the mean man-hours per sortie, but also the variation around this mean. When this is done, it quickly becomes apparent that not all sorties require unscheduled maintenance. Hence two sets of data can be obtained: those that include the "unscheduled maintenance required" information, and those that do not. As will be shown, both sets of measures are important, particularly when dealing with mission-related questions.

It is also apparent that a tabulation of "unscheduled maintenance required" sorties provides the data for computing break-rates, which are the proportion of sorties requiring unscheduled maintenance. These two sets of data, the frequency of breaks and the duration of recovery, provide tremendous leverage on problems of scheduling and operations. Further, since the specific sortie data enable the computation of the variances, the methods of statistical analysis may be used, considerably enhancing the analytic process.

---

*A. Sweetland, The Use of Computers in Air Force Maintenance Management and Analysis, The RAND Corporation, RM-4228-PR, October 1964.

Once we realized how important specific sortie data are, we began developing a computer program (completed in early 1963) to summarize the act of recovery. The program accumulates, for each sortie: the man-hours, the elapsed time (for the recovery), the units produced and break-rate for each set of unscheduled maintenance for the recovery of the entire aircraft, for each system (two-digit work unit code), sub-system (three-digit work unit code), and each work center.*

The data describing the kind of sortie are also included in each record. Thus we have not only the break-rate and recovery data mentioned previously, but also the associated mission data. Hence, we have most of the information necessary to investigate the effects of a modification, i.e., we can determine the effect on the mission, the break- and recovery-rates, and the cost (in units and man-hours).

The missing element in the measurement is the determination of how effective the modification is (unless the objective is to improve the mission, break-rate, recovery or cost figures). For example, many modifications are installed for the comfort and efficiency of the crew. In these instances, it is not only desirable but mandatory to construct, a priori, a set of measures to determine the impact of the modification.

As will be seen, the missing element in the following examples is a set of measures to determine how effective the modification is. Since our purpose is to demonstrate the feasibility of a methodology, this omission is not a serious defect. But for an actual cost-effectiveness analysis, the omission would be intolerable.

The procedures of the mod-costing process divide logically into three parts:

1. The pre-planning phase that includes preparations for the mod-installation. Cardinal among these is the design of a measure to determine the effect of the modification.
2. The installation phase, consisting of
   a. the installation
   b. the checkout-alignment.
3. The post-installation phase that determines the comparative effects of the modification.

---

*Ibid. pp. 4.

We will discuss these three elements beginning with item three, and working backwards -- since this will make the process considerably easier to comprehend. The two real-life exercises already mentioned will be used in illustration: a power plant modification (system 23XXX) done on T-38A aircraft at Williams Air Force Base (ATC), and a fire-control modification (system 744XX) done on F-101B aircraft at Oxnard Air Force Base (ADC).

III. TWO EXAMPLES OF MODIFICATION FIELD TESTING

In our first attempt to develop a method of costing the effects of a modification, we used a sample of Williams Air Force Base data. Since the venture was on an a posteriori basis, there was, necessarily, a good bit of make-do involved. We wanted to know what impact a fairly expensive and extensive power-plant modification had on the T-38A aircraft. The sample was taken from data for the last two weeks of July 1963, when the possessed aircraft count was 84, and the sorties flown totalled 1358.

Inspection showed the sample data could be subdivided into three groups:

1. A control group of unmodified aircraft (CTL).
2. An experimental group modified during the months of May and June (EXF). This group would not be expected to have all the bugs eliminated since the average age of installation was approximately six weeks.
3. An experimental group modified for sufficient time to have stabilized (EXM). These were modified between December 1962 and April 1963. The average age of installation was about three and one-half months.

The data were divided into these groups and then processed by the recovery program to yield the sortie summary data. These latter were analyzed in a number of ways, as described and shown.

The first thing of interest is the break-rate, which identifies (for both system and subsystem) the number of sorties that were followed by unscheduled maintenance. The break-rates for the three aircraft groups are compared in Fig. 1. In this figure, the statistical analysis indicates that, although the "trend" appears to drop from 10/100 to to 8/100, the trend may be attributed to random variation. Even when the test is computed via the more sensitive arc sine method, there is still no justification for assuming that the variation among the measures is other than random.

Following the break-rate analysis, three separate avenues were explored: (1) the analysis of recovery times (system); (2) the analysis

```
ACFT GROUP   SORTIES   BREAKS   MEANS        -3   -2   -1    0    1    2    3    4    5

   EXM         480       39     0.081              .   EXM   .         .         .
   EXF         468       43     0.092              .        EXF        .         .
   CTL         410       42     0.102              .         .   CTL   .         .
         WEIGHTED GRAND MEAN == 0.091                       0.092  (UNWEIGHTED)
         CHI-SQUARE  =  1.20
                 DF  =     2
CHI-SQUARE (ARCSIN)  =  1.20
THE PROBABILITY LEVEL IS BETWEEN 0.70 AND 0.50.  THE NULL HYPOTHESIS CANNOT BE REJECTED.
```

Fig. 1 -- Power Plant (23XXX) Break-rates

(The diagram on the right shows distribution of the means column (39/480 = 0.081). EXM had the modification installed many weeks, EXF had it installed few weeks, and CTL is the control group. The means column shows the break-rate per sortie. Mean and standard deviations of the scores shown in the means column are computed and each score converted to sigma units and displayed graphically at the right. EXM is approximately 1 sigma below the grand mean.)

of man-hour data (system); and (3) the analysis of the recovery of the entire aircraft. Only the first of these, individual system recovery, is discussed in detail here. Examples of points two and three are contained in the Appendix.

The recovery data shown in Fig. 2 were tested in two ways: with and without the inclusion of the "no-break" data. Again, we find a trend in the desired direction. But again we find, as with the break-rate data, that random variation may account for the changes. The data in Fig. 2, incidentally, provide one form of job standard to be used in managing the flight-line (also see Fig. 8).

By adding the no-break data to the data shown in Fig. 2 we determine the impact on the "average" sortie. Combining the no-break and break data, in effect, combines the information of Figs. 1 and 2. These average sortie data are important from the viewpoints of both planning and operations. In terms of capability forecasts, predicted recovery, projected kills, etc., these data show what can be expected "on the average." For example, if we are trying to squeeze training sorties out of our possessed aircraft, these numbers give us planning figures -- expected losses to unscheduled maintenance.

The data of Fig. 1 show a decreasing trend in frequency of unscheduled maintenance, while the data of Fig. 2 show a decreasing trend in the time needed to recover the broken system. The interesting question is whether the two statistically insignificant trends may combine; Fig. 3 shows the results of such efforts. Combining the trends has no effect on the conclusion that only random variation is operating. The differences among the means have no significance.

The findings on the T-38A break-rate and recovery data were corroborated with the man-hour data (Appendix, Figs. 4A and 5A). We also cross-checked the three-digit data (subsystem) with the recovery data of the entire aircraft (Appendix, Figs. 1A to 3A). The results were essentially the same -- a trend in the indicated direction too small to have any statistical significance.

But had there been statistical significance, there would still have been no practical significance. For example, the "savings"

SAMPLE FREQUENCY DISTRIBUTION

WAFB T-38A (RECOVERY-TIME DATA) CODE = (23XXX)

FREQUENCY COUNTS

Z =	-5	-4	-3	-2	-1		0		1		2		3		4		5	MEANS	N	TOTALS
SIGMA	-6.6	-4.9	-3.3	-1.6	0.0		1.7		3.4		5.0		6.7		8.3		10.0			
SAMPLE 1---EXM					4	22	7	2	1	1		1	1					1.49	39	58.00
SAMPLE 2---EXF					3	27	5	2	4							1	1	1.61	43	69.10
SAMPLE 3---CTL					3	18	7	2	6	3	1	1		1				1.98	42	83.30
																	GRAND MEAN	1.70	124	210.40

ANALYSIS OF VARIANCE

	SUM SQUARE	DF	MEAN SQUARE	F
BETWEEN	5.5089	2	2.7545	0.9917
WITHIN	336.0698	121	2.7774	
TOTAL	341.5787	123	2.7771	

Fig. 2 -- Mean Time to Recover the Power Plant (23XXX) Data Shown in Fig. 1

(The data counts in the frequency distribution are clustered into ½-sigma units (although the column headings are given only by unit sigma). Sample 1 -- EXM has 4 measures which are 1 sigma below the mean and 22 measures which are ½-sigma below the mean. The standard deviations are given for each Z score, i.e., +2Z = 5.0, +3Z = 6.7, etc. The totals column shows the hours lost to unscheduled maintenance.)

SAMPLE FREQUENCY DISTRIBUTION

WAFB T-38A (RECOVERY-TIME DATA)

FREQUENCY COUNTS

CODE = 23XXX

Z =	-5	-4	-3	-2	-1	0	1	2	3	4	5
SIGMA	-3.3	-2.6	-1.9	-1.2	-0.5	0.2	0.9	1.6	2.3	3.0	3.7

													MEANS	N	TOTALS
SAMPLE 1---EXM	445	5	15	2	2	5		2		1		1	0.12	480	58.00
SAMPLE 2---EXF	428	5	16	6	4	2	2		2	1	1		0.15	468	69.10
SAMPLE 3---CTL	371	2	14	2	[illegible]	1		2	3	2	1	1	0.20	410	83.30
											GRAND MEAN		0.15	1358	210.40

ANALYSIS OF VARIANCE

	SUM SQUARE	DF	MEAN SQUARE	F
BETWEEN	1.5370	2	0.7685	1.5672
WITHIN	664.4449	1355	0.4904	
TOTAL	665.9818	1357	0.4908	

Fig. 3 -- T-38A Fire-Control Recovery Time

(The no-break data have been combined with the break data shown in Fig. 2.)

resulting from the modification amounts to 0.15 man-hours a sortie. Using an average man-hour cost of $5.54,* the dollar savings amounts to approximately 83 cents a sortie. Even with the fantastic number of sorties that Williams Air Force Base flies (1358 in two weeks), the cost of the modification probably would not be amortized within the life of the system. (With the current flying schedule, the savings on each aircraft are $200 to $300 a year. The cost to install was 275 to 300 man-hours per aircraft, i.e., about $1660.)

We should establish one point firmly: we have been speaking of the impact on the recovery. For the T-38A, recovery was not the primary objective of the modification. The primary objective was to eliminate ground engine stalls, promoting ease of operation. The secondary objective -- to improve the recovery picture and thus increase the number of training sorties -- was, in effect, an afterthought.

The T-38A data has the desirable feature of permitting comparison of experimental and control groups sampled at the same point in time (i.e., the last two weeks in July). The shortcoming of the data is that they do not provide a means of observing the changes over time, i.e., under a variety of stresses. The Oxnard fire-control modification provided the reverse -- absence of a control group from the same point in time, but providing a comparison across time. The advantages and disadvantages are also reversed. The contrast between the two sets of data will highlight a point worth making -- that the really desirable method involves contrasting experimental and control groups across time.

The Oxnard F-101B fire-control modification was designed to improve low-altitude interception capability (primary objective), with no more than a 5-per-cent increase in break-rate and recovery times (secondary objective). As with the previous example, we shall consider only the secondary objective.

The January through May period was a time of manual operation at Oxnard Air Force Base, and thus serves as a base-line. Beginning June 17th the aircraft were sent, serially, to the depot for installation of the modification. On completing the installation, the system was given

---

*AFM 400-12, 28 February 1964, Paragraph 79 F(3). This factor expresses direct man-hours and does not include overhead.

a "smoke test" ("... turn it on and if you do not see or smell anything, send it back") and returned to Oxnard for checkout and alignment. The last aircraft was released (returned to Operations) August 13th. By October, the system had stabilized from the effects of the program. (Although the June-September data are included, only the data from October on should be compared with the January-May base-line when making the critical determination of the impact.)

Figures 4, 5 and 6 show the same information as Figs. 1, 2, and 3 -- the impact on break-rates and recovery times. The data indicate an increase in break and recovery (over the base of January-May) followed by a return to within-bounds condition in January of 1964. These results were, of course, cross-checked with the man-hour and aircraft recovery data (not included). All analyses yield essentially the same information.

To summarize the F-101 analysis, we may tentatively say that the data suggest the modification has had little impact on break- and recovery-rates. But we cannot make a firm statement since we are at a disadvantage in trying to compare data from different points in time. It is said that no month (at an air base) is typical, and herein lies the problem. Oxnard, for example, underwent a complete change in command philosophy in the sample period, going from a centralized to a decentralized control. This might be the the cause of our differences.

Then too, there is a tremendous month-to-month variation in the number of sorties flown, and we have observed a negative correlation between sorties flown (per available aircraft) and break-rates, as well as sorties flown and recovery time. Further, there are variables such as weather, TOC requirements, training requirements, special exercises, fly-bys, etc. that disrupt any even distribution of schedules. Hence, serious doubts arise when comparing data across time because of the problem of differing forces that may be affecting them.

We insist here that the only way to obtain any meaningful information is to have data samples (experimental and control) taken simultaneously, as with T-38A data, under a wide variety of operating conditions, as with the F-101 data. If we cannot have both, we would prefer to have the simultaneous sample. The reasons for this last statement become evident by assuming that the F-101 numbers had shown true and

significant differences: our dilemma then would be to determine whether the differences were due to the modification per se, the varying forces, or a combination of the two.

BASE-LINE	FO	FT	D2/FT	MEANS
332.	99.	115.72	2.42	0.298*
233.	80.	81.22	0.02	0.343
266.	98.	92.72	0.30	0.368
199.	62.	69.37	0.78	0.312
205.	55.	71.46	3.79	0.268**
120.	47.	41.83	0.64	0.392
96.	37.	33.46	0.37	0.385
134.	58.	46.71	2.73	0.433*
213.	83.	74.25	1.03	0.390
314.	115.	109.45	0.28	0.366
169.	70.	58.91	2.09	0.414
235.	73.	81.91	0.97	0.311
		WEIGHTED GRAND MEAN =		0.349
		CHI-SQUARE =	23.68	
		DF =	11	

```
-3  -2  -1   0   1   2   3   4   5
    .  01   .       .       .
    .      02       .       .
    .       03      .       .
    .   04  .       .       .
    05      .       .       .
    .       . 06    .       .
    .       . 07    .       .
    .       .    08 .       .
    .       . 09    .       .
    .       10      .       .
    .       .   11  .       .
    .   01  .       .       .
            0.357 (UNWEIGHTED)
```

THE PROBABILITY LEVEL IS BETWEEN 0.02 AND 0.01 THE DISTRIBUTION IS MOST LIKELY, NOT HOMOGENEOUS.

CHI-SQUARE (ARCSIN) = 23.91

THE PROBABILITY LEVEL IS BETWEEN 0.02 AND 0.01 THE DISTRIBUTION IS, MOST LIKELY, NOT HOMOGENEOUS.

Fig. 4 -- F-101 Break-rates for Fire Control

(This figure is similar to Fig. 1 but includes more information. The base-line gives the sorties-flown count, FO is the observed frequency of breaks. The theoretical frequencies (FT), and the $(FO-FT)^2/FT$ entries are part of the conventional chi-square computations. They are included to facilitate analysis. The means column (FO/base-line) has asterisk tags beside those items that show significant differences between the FO and FT columns.)

SAMPLE FREQUENCY DISTRIBUTION

F-101 RECOVERY (ELAPSED-TIME DATA) SYSTEM 744XX INTERCEPT SORTIES ONLY   CODE = 744

FREQUENCY COUNTS

Z =	-5	-4	-3	-2	-1		0		1		2		3		4		5	MEANS	N	TOTALS
SIGMA	-24.2	-18.4	-12.6	-6.8	-1.0		4.8		10.7		16.5		22.3		28.1		33.9			
SAMPLE 1---01					2	53	25	7	6	2	1	2	1					4.25	99	421.00
SAMPLE 2---02					1	41	17	7	6	5		2				1		5.00	80	399.80
SAMPLE 3---03						43	21	12	9	4	1	1	3	2	2			6.12	98	599.50
SAMPLE 4---04						30	19	5	2	1	2	2		1				5.04	62	312.60
SAMPLE 5---05					4	32	7	7	3	2								3.48	55	191.40
SAMPLE 6---06						26	9	5	3			1			1			6.03	47	283.40
SAMPLE 7---07						15	12	1	3	1		2	1					8.26***	37	305.60
SAMPLE 8---08					3	29	15	6	1	1		2		1				4.14	58	240.30
SAMPLE 9---09					2	42	15	13	7	1		1	1				1	4.79	83	397.40
SAMPLE 10---10					7	57	31	8	7	2	2		1					3.96	115	455.60
SAMPLE 11---11					1	31	18	11	4			1	2		1	1		5.61	70	392.90
SAMPLE 12---01					3	47	12	6		3	2							3.29*	73	240.00
														GRAND MEAN				4.83	877	4239.50

ANALYSIS OF VARIANCE

	SUM SQUARE	DF	MEAN SQUARE	F
BETWEEN	1134.2822	11	103.1166	3.1258
WITHIN	28535.5747	865	32.9891	
TOTAL	29669.8569	876	33.8697	

Fig. 5 -- F-101 Fire-Control Recovery Times

(January 1963 -- Sample 1, Month 01; through January 1964 -- Sample 12, Month 01. The F-Test is highly significant. Note that the October mean is well within the January-May range. The no-break data are not included. As in Fig. 2, frequency counts are clustered at the ½-sigma points. Items contributing most to the F-Test are tagged with asterisks.)

SAMPLE FREQUENCY DISTRIBUTION

F-101 RECOVERY (ELAPSED-TIME DATA) SYSTEM 744XX INTERCEPT SORTIES ONLY CODE = 744

FREQUENCY COUNTS

Z =	-5	-4	-3	-2	-1	0	1	2	3	4	5
SIGMA	-19.0	-14.9	-10.7	-6.6	-2.5	1.7	5.8	10.0	14.1	18.2	22.4

														MEANS	N	TOTALS
SAMPLE 1---01	240	43	17	15	5	4	3	1	1	1	1	1		1.27	332	421.00
SAMPLE 2---02	157	35	13	9	4	4	5	3		2				1.72	233	399.80
SAMPLE 3---03	174	34	19	10	7	5	6	2		2		3	2	2.25	266	599.50
SAMPLE 4---04	138	24	15	9	5	1	1	1	1	3				1.57	199	312.60
SAMPLE 5---05	160	21	9	4	5	4	2							0.93*	205	191.40
SAMPLE 6---06	77	20	5	10	1		3				1			2.36*	120	283.40
SAMPLE 7---07	62	10	6	8	1	2	2			1	1		1	3.18***	96	305.60
SAMPLE 8---08	85	20	13	6	4	2	1			1	1		1	1.79	134	240.30
SAMPLE 9---09	138	33	13	7	11	4	3	1			1	1		1.87	213	397.40
SAMPLE 10---10	214	45	21	16	6	3	5	1	2			1		1.45	314	455.60
SAMPLE 11---11	102	25	12	12	9	2	2				1	2		2.32*	169	392.90
SAMPLE 12---01	178	30	11	5	6		2	1	1	1				1.02*	235	240.00
GRAND MEAN														1.69	2516	4239.50

ANALYSIS OF VARIANCE

	SUM SQUARE	DF	MEAN SQUARE	F
BETWEEN	731.2385	11	66.4762	3.9362
WITHIN	42289.1416	2504	16.8886	
TOTAL	43020.3804	2515	17.1055	

Fig. 6 -- F-101 Fire-Control Recovery Times

(These are the break data shown in Fig. 5 with the no-break sorties added. The F-Test is highly significant.)

IV. MONITORING INSTALLATION AND CHECKOUT

To facilitate the fire-control checkout at other bases, the 28th Air Division requested that Oxnard provide modification data based on the first 12 aircraft it released back to Operations. The list of Oxnard contributions included among other things:

1. The "queue sorts" on each aircraft -- a pictorial history of the alignment and checkout procedures
2. The elapsed time for completing the checkout process for each aircraft
3. The man-hours and units produced for each aircraft
4. The total bit and piece consumption for all 12 aircraft
5. The number of test-flights necessary to learn whether the new equipment was functioning satisfactorily
6. The kinds of defects discovered in the new system (i.e., flight-crew write-ups)
7. The revised job standards.

Many of these items were satisfied by a rather peculiar kind of listing. In the beginning, we created an output that described the checkout process almost entirely in English. One of the computer programs we had developed provided for the inclusion of in-English statements that were key-punched and tagged with a unique card-code. In producing the daily monitoring program, the computer, on sensing this code, printed the comment verbatim.

In this new output, we instructed the computer to print the verbal statement along with the accumulated units and man-hour data for each aircraft. The result was, essentially, a word-history of the checkout process. A page of this product is shown in Fig. 7, which gives the highlights of the checkout for three aircraft -- tail numbers 7434, 7436 and 7438.

The time entries locate each event. The elapsed time of installation is obtained by subtracting time begun from time end, and, where necessary, adjusting for NORS condition. The flight crew write-ups are shown for each sortie ("not released" indicating not released back to operations.) The units produced and man-hour totals conclude the summary for each aircraft.

HUGHES MOD 1077 OXNARD IIP FIRST 12 A/C---TOUCH-DOWN TO RELEASF

TIME-DISC DAMO	HRMN		UNITS	TOT MAN -HOURS-	SERIAL NUMBER
2806	1500	RETURNED FROM DEPOT 1500 HOURS			7434FF
2806	1600	STARTED WORK 1600 HOURS			
807	800	COMPLETED 0800 HOURS			
807	801	NORS (STRUT AND SEAL)			
907	1	NORS (STRUT AND SEAL)			
1007	1	NORS (STRUT AND SEAL)			
1107	920	NOT RELEASED---WEAK PICKUP-IR STEERING-DOOR ROTATED EARLY			
1307	1350	NOT RELEASED---WEAK PICKUP RADAR AND IR			
1407	915	NOT RELEASED---GENERATOR FAILED-BAD MSR			
1507	910	NOT RELEASED---CHANGING RANGE SELECTION BREAKS LOCK			
1507	911	IR STEERS LEFT OF TARGET			
1607	1620	RELEASED			
			87	421.6•	
507	1600	RETURNED FROM DEPOT 1600 HOURS			7436FF
707	800	STARTED WORK 0800 HOURS			
1107	1000	COMPLETED 1000 HOURS			
1207	1	NORS (LANDING GEAR ACTUATOR)			
1307	1	NORS (LANDING GEAR ACTUATOR)			
1407	1	NORS (LANDING GEAR ACTUATOR)			
1507	1	NORS (LANDING GEAR ACTUATOR)			
1707	930	NOT RELEASED---LOST ALL PRESENTATIONS			
1807	1255	NOT RELEASED---LOST VERTICAL DEFLECTION BOTH SCOPES			
1907	915	NOT RELEASED---SCOPES WENT FLAT-MSR TIME OF FLIGHT WRONG			
2107	920	NOT RELEASED---NO RANGE GATE MARKER AFTER 1ST COINCIDENCE			
2307	910	NOT RELEASED---ABORTS 2 SEC AFTER LOCKON-STEERING ERRATIC			
2507	910	RELEASED			
			89	334.2•	
607	1300	RETURNED FROM DEPOT 1300 HOURS			7438FF
907	900	STARTED WORK 0900 HOURS			
1607	1900	COMPLETED 1900 HOURS			
1607	2000	NORS (AILERON ACTUATOR)			
1707	1	NORS (AILERON ACTUATOR)			
1907	910	NOT RELEASED---LOST POWER SUPPLY ON TAKEOFF			
2107	900	GROUND ABORT---RADAR HORIZON TILTED 30 DEGREES			
2107	901	UNABLE TO TUNE IR AURAL NULL			
2107	1340	NOT RELEASED---SCOPES ARCING ABOVE 40,000 - IMPROPER			
2107	1341	COURSE IR DOMINANT STEERING			
2307	910	NOT RELEASED---IR STEERING BAD-INTERMITTENT LOSS OF VIDEO			
2607	915	RELEASED			
			107	384.5•	

Fig. 7 -- Verbal Records

(This figure shows the history during the alignment and check-out of the fire control modification. DAMO HRMN represents Day, Month, Hour and Minute.)

Job standards were computed for each of the five-digit work-unit codes in subsystem 744XX. Only one of these is shown, the amplifier relay assembly, work-unit code 7444E. Figure 8 depicts the job standards program output -- two juxtaposed histograms identical except for each of the cell entries. Each similarly located pair of cells represents one 200 series document. The program produces an approximation of job duration under a wide variety of actions taken. Since these data are inevitably of the-standard-deviation-equals-the-mean variety (i.e., no possibility of predicting their duration), the probability of being completed by various points in time is computed (the "P-" line under the left histogram.)

Figure 8 well illustrates the capability of the job standards program to winnow information from a limited amount of data. In it, we see summarized the entire set of actions on work unit 7444E. (Incidentally, the isolated "5" at 6 hours is undoubtedly a keypunch error since "5" is the code for "deferred.")

The reader has undoubtedly noticed a lacunae in both samples: the installation data (i.e., the AFTO 212 information). In neither instance was this information available. The F-101 modification was installed at the depot. The T-38A data were not available (i.e., occurred before the inclusion of clock-hours in the AFM 66-1 data). No difficulty is anticipated in processing the 212 data, since the only statistical accumulation is a man-hour total for each unit produced (i.e., installation). The one addition to the regular 66-1 system is the determination of installation elapsed times. These can be obtained by any of four methods: using the verbal inputs as in the F-101 data, manually checking the daily monitoring products, or making a special pass through the recovery program. A less elegant, but adequate, method is to divide total man-hours by average team size.

```
                              FIVE-DIGIT SYSTEM  7444E

                 TEAM SIZE                                               ACTION TAKEN

 40                                                      40
 39                                                      39
 38                                                      38
 37                                                      37
 36                                                      36
 35                                                      35
 34                                                      34
 33                                                      33
 32                                                      32
 31                                                      31
 30                                                      30
 29                                                      29
 28                                                      28
 27     2                                                27     S
 26     2                                                26     C
 25     1                                                25     C
 24     2                                                24     F
 23     2                                                23     F
 22     2                                                22     S
 21     2                                                21     S
 20     2                                                20     C
 19     1                                                19     F
 18     2                                                18     S
 17     2                                                17     S
 16     3                                                16     S
 15     1 1                                              15     S S
 14     1 2                                              14     F S
 13    22 2                                              13    QF Q
 12    22 2                                              12    CC S
 11    22 2                                              11    LS R
 10    12 2                                              10    PC R
  9    22 2                                               9    PS S
  8    22 1                                               8    PR S
  7    12 2 2                                             7    PS R S
  6    12 3 2                                             6    PS R S
  5    2121 2                                             5    PCRP R
  4    2122 222 2                                         4    KFPK RSY S
  3    22122222 2 1                                       3    KKSKQRRS R R
  2    12221322 2 2                                       2    PSSSSSSG S S
  1    22222422 2223  22      1                           1    PPYSSLSS SSSS  YS     5

HOURS  *  1   2   3   4   5   6   7   8   9  10  11  12    HOURS  *  1  2  3  4  5  6  7  8  9  10  11  12
P =      66  86  95  97  98 10C 10C 10C 100 10C 1C0 10C

    AVG =     1.16       SIGMA =  1.02      N =  9C      STDERR =  0.11      VARIANCE =  1.05      TOTAL =     104.00
 TOTAL MAN HOURS =      99.7 SATURATION INDEX =  1.0     UNITS PRODUCED =   89
```

Fig. 8 -- Job Standards for the Amplified Automatic Relay

(The P, Q, R, and S codes represent removal and replacement actions. The "5" is undoubtedly a key-punch error.)

V. FUTURE DEVELOPMENTS

At first glance, it would seem that the methods suggested above could be used in pretesting a modification to determine whether to purchase. Realistically, such a program is not feasible. The time it takes to develop a modification may be (and often is) measured in years. This fact, coupled with the expense, precludes any "back to the drawing board" modus operandi. Instead, modifications are bought as "specs" i.e., the fire control was bought as a specification "to improve low-altitude intercept capability..." (capability is defined in detail). The particular way the manufacturers will meet these specifications is immaterial at this point.

Developing hardware to meet the specification is often costly: indeed, development costs may be several times as great as the final hardware costs. Hence, the interest in determining modification impact is not to accept or reject a product, but rather to determine whether the specification is met. If not, it is reasonable to expect a delay of acceptance. In general, delay does not occur since specifications were bought rather than hardware.

The previous paragraph is not meant to imply that the process is uneventful. Quite the contrary. It is inevitably difficult and often impossible to establish reliable and valid operational definitions on an a priori basis. Even the simple T-38A specification, "...the elimination of J-85 engine stalls, ground," demands some thought. How long must an engine have run before a stoppage can be called a stall, two seconds? ten seconds? ten minutes? We have all had the experience of having an automobile cough twice and die. If an aircraft does this, does it constitute a stall? If immediate restart is possible, is the stall as significant as a stall that occurred when taxiing (i.e., disconnected from ground support equipment). The questions appeared trivial until we realized that about $17 million was involved.

Despite these apparent difficulties, it does seem possible to define a set of terms that are reasonably unequivocal. And, once these are had, to record the number of engine stalls under various

conditions: taxi, test-cell, run-up, take-off, etc. Also, since the overwhelming majority of tech-orders are concerned with system or component reliability, the possibility of obtaining by-sortie data implies that before-and-after comparisons are feasible, and further, that we can statistically determine whether the differences are real or are random artifacts. The need for adequate control groups (as in the T-38A engine analysis) is re-emphasized.

Measuring the improvement in direct mission effectiveness is, of course, impossible, since the only true measures are those taken from actual combat, and we have no desire to validate such a set of measures. Instead, we shall have to be content to infer the results from measures taken in peacetime conditions. Thus, it seems reasonable to conclude that an interception modification that reduces peacetime break- and recovery-rates would also improve combat kill probabilities. We would like, however, some more direct measures of mission accomplishment: how good was the bomb drop? Would the air-to-air missile have killed the target? At present, at least two such measures show promise: the SAC U-82 reports for aircraft, GAMS, and missiles (see SACR 66-7), and the ADC 76-3 reports for intercept success (see ADCR 66-28). Both documents are particularly admirable because they permit retrieval of "total-items-attempted-and-total-items-successful" information. Hence, the data are amenable to statistical methods. In addition, the ADC document describes in detail the circumstances and characteristics of both target and interceptor at the time of intercept attempt. Thus, one can determine what kinds of tactics yield the best kill ratio for a wide variety of circumstances.

Both the SAC and ADC reports have a serious deficiency, however: it is nearly impossible to relate the data to the associated maintenance. Without special techniques, one cannot determine, for example, what methods of maintenance result in degraded mission performance, or what material failures are specifically responsible for the loss of the mission. Note that the latter is obtainable from the narrative accounts, but in no instance, is this information available to the Logistics Command.

Hence, there is need for prolonged exploration to relate the mission and maintenance data and, thereby, to assess the impact of material and maintenance on the mission, and the inverse impact of the mission on material maintenance. This exploration should result in the development of better methods for determining modification impact and, more important, the suggestion of modifications that would improve mission effectiveness.

There is another area that needs considerable exploration: the determination of those factors affecting the break- and recovery-rates. For example, it is well-known among ADC maintenance people that high-altitude (over 50,000 feet) intercepts result in increased power plant and radar problems. Current unpublished studies show a score of such factors operating, the most predominant one being the human decision factor. There is a need for a complete exploration to determine what factors are the dominant affecters of break- and recovery-rates, in order to make more meaningful management decisions such as "what resources to preposition under what circumstances."

A third area to explore, test design, can be expected to make a substantial contribution toward creating more adequate methods of measurement. Over the years, a large number of mathematical and statistical methods have been developed to enable the construction of a variety of tests (such as psychological testing for I.Q., emotional stability, vocational interest, and vocational aptitude; or industrial testing for process, control, acceptance inspection, and reliability.) The more we work with the problem of maintenance analysis, the more we are realizing that we have available a large collection of rigorous methods for helping us evaluate the effectiveness of maintenance, material, and operations. Some of the more recent developments in multi-variate techniques appear to be particularly promising because they facilitate the evaluation of the effects of simultaneous impact of a wide variety of elements.

VI. CONCLUSIONS

Given some comparatively minor additions to current Air Force documentary procedures, major increases in quality and depth of analysis of maintenance material problems are possible. In this Memorandum we have shown that the ability to isolate information by the unique sortie provides break- and recovery-rate data not previously available. The usefulness of these data was demonstrated by two cases -- a T-38A engine modification and an F-101 radar modification. It was further concluded that mission data must be available in a form that can be related to maintenance data in order to make a complete picture.

APPENDIX

This Appendix contains Williams Air Force Base T-38A analyses. Figures 1A to 3A correspond to Figs. 1 to 3 in the text. The data show the recovery of the entire aircraft. Figures 4A and 5A show the man-hours used to correct power-plant failures.

Similar tests were also made with the units-produced data. It should be noted that there are high, positive correlations among these measures (recovery times, man-hours, and units produced). Hence they cannot be viewed as independent evaluations.

```
            WAFB T-38A   AIRCRAFT RECOVERY TIMES

  BASE-LINE     FO        FT       D2/FT       MEANS               -3  -2  -1   0   1   2   3   4   5
     468.      151.    161.40      0.57        0.323                   . EXF    .       .       .
     480.      168.    165.54      0.04        0.350                   .        EXP     .       .
     409.      143.    141.06      0.45        0.364                   .        .  CTL  .       .
            WEIGHTED GRAND MEAN =              0.345                          0.346 (UNWEIGHTED)
           CHI-SQUARE =        1.76
                   DF =           2
THE PROBABILITY LEVEL IS BETWEEN 0.50 AND 0.30  THE NULL HYPOTHESIS CANNOT BE REJECTED.
CHI-SQUARE (ARCSIN) =      1.77
THE PROBABILITY LEVEL IS BETWEEN 0.50 AND 0.30  THE NULL HYPOTHESIS CANNOT BE REJECTED.
```

Fig. 1A -- T-38A Break-rates Computed for the Entire Aircraft (All Systems Included)

SAMPLE FREQUENCY DISTRIBUTION

WAFB T-38A AIRCRAFT RECOVERY TIMES CODE = A/C

FREQUENCY COUNTS

Z =	-5	-4	-3	-2	-1	0	1	2	3	4	5
SIGMA	-26.8	20.8	-14.8	-8.8	-2.8	3.2	9.2	15.7	21.2	27.2	33.2

	FREQUENCY COUNTS	MEANS	N	TOTALS
SAMPLE 1---EXF	56 61 25 7 2	3.00	151	452.70
SAMPLE 2---EXP	60 75 21 9 1 1	3.67	168	616.60
SAMPLE 3---CTL	63 60 18 6 2	2.86	149	426.10
GRAND MEAN		3.20	468	1495.40

ANALYSIS OF VARIANCE

	SUM SQUARE	DF	MEAN SQUARE	F
BETWEEN	60.5560	2	30.2780	0.8406
WITHIN	16749.5979	465	36.0206	
TOTAL	16810.1541	467	35.9960	

Fig. 2A -- T-38A Recovery Times

(Shows mean time used to recover the entire aircraft.)

SAMPLE FREQUENCY DISTRIBUTION

WAFB T-38A AIRCRAFT RECOVERY TIMES CODE = A/C

FREQUENCY COUNTS

Z =	-5	-4	-3	-2	-1	0	1	2	3	4	5
SIGMA	-18.1	-14.2	-10.4	-6.6	-2.7	1.1	4.9	8.H	12.6	16.4	20.3

	FREQUENCY COUNTS	MEANS	N	TOTALS
SAMPLE 1---EXF	317 67 41 24 10 7 1 1	0.97	468	452.70
SAMPLE 2---EXM	312 80 45 26 7 6 1 1 1	1.28	480	616.60
SAMPLE 3---CTL	260 74 41 17 9 3 4 1	1.04	409	426.10
GRAND MEAN		1.10	1357	1495.40

ANALYSIS OF VARIANCE

	SUM SQUARE	DF	MEAN SQUARE	F
BETWEEN	25.9818	2	12.9909	0.8833
WITHIN	19914.5059	1354	14.7079	
TOTAL	19940.4878	1356	14.7054	

Fig. 3A -- T-38A Recovery Rates Averaged Across All Sorties

SAMPLE FREQUENCY DISTRIBUTION

WAFB T-38A (MAN-HOUR DATA)

CODE = 23

FREQUENCY COUNTS

Z =	-5	-4	-3	-2	-1	0	1	2	3	4	5
SIGMA	-9.9	-7.3	-4.7	-2.0	0.6	3.2	5.9	8.5	11.1	13.8	16.4

												MEANS	N	TOTALS
SAMPLE 1---EXM	10	11	2	11	2		2		1			3.05	39	119.00
SAMPLE 2---EXF	10	19	1	9	2			2				2.83	43	121.60
SAMPLE 3---CTL	8	11	8	4	6	1		1	1	1	1	3.82	42	160.40
GRAND MEAN												3.23	124	401.00

ANALYSIS OF VARIANCE

	SUM SQUARE	DF	MEAN SQUARE	F
BETWEEN	22.7692	2	11.3846	1.6500
WITHIN	834.8486	121	6.8996	
TOTAL	857.6178	123	6.9725	

Fig. 4A -- T-38A Average Man-hours Used to Fix Power-Plant Failures

(Break-data only. Because of the high correlation between recovery time and man-hours, these data contribute little additional information.)

SAMPLE FREQUENCY DISTRIBUTION

WAFB T-38A (MAN-HOUR DATA) CODE = 23

FREQUENCY COUNTS

Z =	-5	-4	-3	-2	-1	0	1	2	3	4	5
SIGMA	-5.8	-4.6	-3.4	-2.2	-0.9	0.3	1.5	2.7	4.0	5.2	6.4

												MEANS	N	TOTALS
SAMPLE 1---EXM	449	2	1	8	4		9		2	2		0.25	480	119.00
SAMPLE 2---EXF	428	7	4	10	5		4	4	2	2		0.26	468	121.60
SAMPLE 3---CTL	373	3	1	7	9	3	2	1	1	5	2	0.39	410	160.40
GRAND MEAN												0.30	1358	401.00

ANALYSIS OF VARIANCE

	SUM SQUARE	DF	MEAN SQUARE	F
BETWEEN	5.4387	2	2.7194	1.8146
WITHIN	2030.5510	1355	1.4986	
TOTAL	2035.9897	1357	1.5004	

Fig. 5A -- T-38A Average Man-hours to Fix Power-Plant Failures (All Sorties)